GROUP-TESTING, HALVING PROCEDURES AND BINARY SEARCH

by

Bernard Pasternack[*1], Milton Sobel[**2], and James Thomas[***]

Technical Report No. 227

March, 1974

[*] Dept. of Environmental Medicine, NYU Medical Center, New York, NY 10016

[**] Dept. of Theoretical Statistics, University of Minnesota, Mpls., Mn. 55455

[***] Dept. of Radiology, SUNY Downstate Medical Center, Brooklyn, NY 11203

[1] Supported by Grant No. ES-00260 from the National Institute of Environmental Health Sciences at New York University Medical Center.

[2] Supported by U.S. Army Grant DA-ARO-D-31-124-72-G187 at the University of Minnesota, Minneapolis, Minnesota.

Introduction.

Group-testing is concerned with the classification of N units into disjoint categories. In the simplest case there are two distinct categories which are called good and defective. The characteristic feature of unrestricted binary group-testing is that any number x of units $(1 \le x \le N)$ can be tested simultaneously with only one of two possible error-free outcomes: (i) all the x are good or (ii) at least one unit among the x is defective (it is not known which ones or how many). The problem is to devise a scheme, preferably sequential, which minimizes the expected number of tests needed to classify all of the N units as either good or defective. The units are assumed to have come independently from a binomial population with common probability p of being defective and $q = 1-p$ of being good.

In this paper it is assumed that q is known. The case of unknown q is discussed in [7] and [9]; in the latter a Bayes solution is found for a beta prior with small integer exponents. An asymptotic Bayes solution is developed in [8].

There are numerous versions of this problem depending on exactly what is known beforehand, how many and what outcomes each test on x units can have, and different restrictions that can be placed on the above formulation; some different versions can be found among the references at the end of this paper. Many papers have been written on group-testing since the first paper by Dorfman [1]; his procedure is described near the end of this paper.

Perhaps the simplest version of the problem (now almost classical) is Problem one in which it is known at the outset that there is exactly one defective present in the whole set of N units. In this version the value of q has no bearing on the problem. It has been shown by Huffman [2], Sandelius [6], and Zimmerman [18], that for this case the optimal solution is a halving procedure R_{HSZ} that clearly does not start by testing the entire set of N units. If $N = 2^y + R$ where y and R are nonnegative integers with $R < 2^y$, then in the first test any subgroup of size x is used, where x is in the closed interval $[2^{y-1}, 2^y]$ and eliminate whichever group contains no defective. To be more specific, for the halving procedure, it can be assumed that $0 \leq N/2 - x < 1$ or that $x = [N/2]$. Letting n denote the current size of the defective batch, i.e., the batch that contains the defective units, this procedure is repeated with n replacing N until the defective unit is identified. The expected number of tests needed, $E(T)$, is known (e.g., see [6]) to be

$$E\{T|R_{HSZ}\} = y + \frac{2R}{N} .$$

Thus, roughly speaking, a halving procedure is one that uses one of the integers closest to $n/2$ whenever it is necessary to select a subset of units from these n units for testing. Later on in the paper a procedure is regarded as a halving procedure even though it uses halving only when a defective batch is partitioned. The idea of binary search refers to the fact that there are only two disjoint outcomes for each test, regardless of how many units are being tested.

The case in which it is known beforehand that there are exactly $D \geq 2$ defectives present also does not depend on q, but the optimal solution for locating all the defectives is <u>not</u> a halving procedure. A nested procedure based on recursive formulas for this problem is developed in [12]; for $D = 2$ it is shown to be optimal among all procedures for an infinite set of possible starting N-values.

Since the halving procedure above is optimal in the classical problem with exactly 1 defective and was also noted to be asymptotically ($q \to 1$) optimal for fixed N in other problems (e.g., see [7] and also remarks below for the problem of finding exactly 1 defective), it is quite natural for people to think of using this procedure for all kinds of group-testing problems. Hence it is of considerable importance to know just when some halving procedure is optimal and to know its limitations in various problems. It is with these ideas in mind that the present paper is being written, i.e., we wish to compare certain halving procedures with the corresponding optimal procedures for: (i) the problem of finding one defective if it exists, and (ii) the problem of finding all the defectives.

All the procedures considered in this paper have a nested property in the sense that when we have a batch that contains a defective unit, one continues to break down, i.e., partition, that batch until a defective unit is found. Moreover, this is completed without mixing in any other units from other sources. This property of being nested is inherent in the recursive formulas that define the procedure.

Problem Two: Finding One Defective if it Exists.

In a recent paper, Thomas, Pasternack, Vocirca and Thompson [16] introduced a procedure R_T for application to a problem of radiological health, in which the entire batch is tested at the outset and, if the units are not all good, then we proceed to look for exactly 1 defective unit using the halving procedure mentioned above. (It is not essential to test the entire set of N units on the first test; this helps to keep the expected number of tests small when q is close to 1, but does not help to increase the probability, $P(CC)$, of correctly classifying all units, since all procedures that find one defective (whenever one exists) have the same $P(CC)$ given by (8) below.) Using n to denote the current size of the defective batch (as above), the next test batch of size x is taken (from the defective batch) to be such that $0 \le n/2 - x < 1$, i.e., $x = [n/2]$ is taken in a nested manner from the batch that is known to contain at least one defective.

The expected number of test $E\{T|R_T\}$ for this procedure also denoted by $H_T(N)$ or simply by $H(N)$, can be easily obtained from the following recursions:

For $N \ge 1$

$$H(N) = 1 + p\, F^*(N) \; . \tag{2}$$

For $n \ge 2$

$$F^*(n) = \frac{1-q^n}{1-q} + q^x F^*(n-x) + F^*(x) \; , \tag{3}$$

where $x = [n/2]$. The boundary conditions for this recursion are $H(0) = F^*(1) = 0$. Here $F^*(x) = p^{-1}(1-q^x)\, F(x)$ where $F(x)$ is the expected number of tests needed to find a defective in a batch of size x that is known to contain at least one defective. The derivation of (2) and (3) is along the same lines as used in [3], [7], and [10] and need not be repeated here. Equations (2) and (3) can be combined to form a single recursion of the form

(4) $$H(N) = 1 - q^x - q^N + q^x H(N-x) + H(x) \qquad \text{for} \quad N \geq 2,$$

where $x = [N/2]$ and only the single boundary condition, $H(1) = 1$ is needed. Thus, e.g., for $N = 15, 30, 50$, we obtain after a fair amount of algebra

(5) $$H_T(15) = 4(1-q^{15}) + q \ ,$$

(6) $$H_T(30) = 5(1-q^{30}) + q - pq^{15} \ ,$$

(7) $$H_T(50) = 6(1-q^{50}) + q - \frac{(1 - q^{18} + q^{22} - q^{43})}{1 + q + q^2} \ ;$$

the numerical values for $q = .90$ in the above are 4.0764, 5.6675 and 6.5231, respectively.

We are also interested in the probability that the use of procedure R_T will lead to the correct classification of all units, i.e., that either all units are good or the defective unit found is the only defective unit present. In other words, all units not found to be defective are tacitly assumed to be good under procedure R_T. For any $N \geq 1$ and any q this probability $P(CC)$ is clearly

(8) $$P\{CC|R_T\} = q^N + Npq^{N-1} \ ,$$

and the numerical values for $q = .9$ in the above 3 cases are .5490, .1837, and .0338, respectively. Clearly, if $q \to 1$ then $P(CC) \to 1$ for any fixed N but for fixed values of q and large values of N the result can be quite far from 1 as is seen above. Thys the procedure R_T should not be used for the purpose of finding all the defectives with high probability unless n is moderate in size and p is very close to zero.

It is desirable to avoid making comparisions of the $E\{T\}$ for different procedures unless the $P(CC)$-values are exactly the same. The $P(CC)$-value is the usual (frequency-type) probability and refers to what can be expected if the procedure is used over and over again; it is not to be confused with the posterior probability of a correct classification of all units, which depends strongly on what is observed for the particular sample at hand.

Three different lower bounds for $H_T(N)$ will be discussed. One is from coding theory and is called the Huffman lower bound (HLB) (See [2]). Another from information theory has a simple analytic form and is called the information lower bound (ILB). (See [7]). Both of these bounds apply to any group-testing procedure that finds at most one defective unit and hence they enable the user to assess how close any procedure for the problem at hand is to being an optimal procedure. The third bound is actually a pair of bounds (upper and lower), but they apply only to the procedure R_T. Both the HLB and ILB are based on the identity

$$(9) \qquad q^N + p + qp + \ldots + q^{N-1} p = 1 ,$$

which gives the probabilities of the $N + 1$ possible outcomes of any procedure applicable for Problem 2; here qp (for example) indicates that the defective unit found was in the second position (after the N units were initially randomized and the resulting order held fixed) and the unit in the first position was found to be good.

For the HLB these $N + 1$ probabilities are ordered and the 2 smallest are added, thus forming the first new number which replaces the two that gave rise to it. This resulting set of N numbers is reordered and again the two smallest are added, forming the second new number. This process is repeated until only one new number, whose value is one, remains. The sum of all the N new numbers is the BLB. No simple explicit analytic expression holds for the HLB, but the calculation for any q and moderate N is straightforward and easy with or without a computer. For $N = 15$ and $q = .90$ the value obtained is

$$\text{HLB} = 3.7606 \ . \tag{10}$$

The ILB is obtained by the formula

$$\text{ILB} = -\sum_{i=0}^{N} p_i \log_2 p_i \tag{11}$$

where the p_i are the $N + 1$ probabilities on the left side of (9). A little algebra gives the result

$$\text{ILB} = -\left(\frac{1-q^N}{1-q}\right)(p \log_2 p + q \log_2 q) \ . \tag{12}$$

This is in agreement with the result in (7.4) of [3] for the case $N = \infty$. For $N = 15$ and $q = .90$ the value of (12) is

$$\text{ILB} = 3.7244 \ . \tag{13}$$

For the third bound let the random variable $\mathcal{D}$ denote the (true) number of defectives among the N units. Then

$$\begin{aligned} H(N) &= \sum_{d=0}^{N} P\{\mathcal{D} = d|N\}\ E\{T|\mathcal{D} = d\} \\ &\leq q^N + (1-q^N)\ E\{T|\mathcal{D} = 1\} \end{aligned} \tag{14}$$

since $\mathcal{D} > 1$ leads to an easier problem than $\mathcal{D} = 1$. Using (1) and the fact that all the N units are tested at the outset, one obtains

$$E\{T|\mathcal{D} = 1\} = y + 1 + \frac{2R}{N} , \tag{15}$$

where y and R are defined above by writing $N = 2^y + R$. The upper bound, obtained by putting (15) in (14), is

$$H(N) \le q^N + (1-q^N)(y + 1 + \frac{2R}{N}) = 1 + (1-q^N)(y + \frac{2R}{N}) . \tag{16}$$

Moreover, since $y + 1$ is the minimum number of tests for $\mathcal{D} \ge 1$, one also has the lower bound

$$H(N) \ge q^N + (1-q^N)(y + 1) = 1 + y(1-q^N) . \tag{17}$$

For $N = 15$ and $q = .90$ these yield the bounds

$$3.3823 \le H(15) \le 4.1235 \tag{18}$$

but it should be noted that the procedure R_T was used in the derivation. It should also be noted that the classical halving procedure R_0 that starts by testing half of the N units (i.e., by $x = [N/2]$) and makes an extra final test whenever necessary accomplishes the same task, namely, it finds a defective unit if one exists and stops as soon as it is found. However, the analysis now depends on q and the simple result (1) no longer holds, not even as a bound (except possibly for $N = 2^y$ in which case

$$E(T|R_0) = y + q^{N-1} \ge y \tag{19}$$

and the lower bound y is also obtained from (1) when $R = 0$). However, for q close to 1 the halving procedure R_T is better than R_0 since the chance of stopping after 1 test under R_T is not negligible. For $N = 15$ it is easily shown that

$$(20) \qquad E(T|R_0) = 3 + q + q^{14}$$

and for $q > q_0$ (where $q_0 < .9$) the value in (5) is less than that in (20). In particular, $N = 15$ and $q = .90$ it has already been seen that $H_T(N) = 4.0764$ and by (20) $E(T|R_0) = 4.1288$. A generalized halving procedure should allow us to use R_0 or R_T, whichever gives the better result. The P(CC) in (8) holds for all of these procedures.

The authors of [15] suggest the use of procedure R_T as a means of classifying all the units, by assuming that for certain applications (e.g., testing sealed radium sources for leakage as in [16]) the probability of more than one defective being present is negligible; here units not found to be defective are classified as good. They recognized the uncertainty of this procedure (the P(CC) is given by (8)), but did not explicitly examine the numerical value of this uncertainty or its implications. Subsequently Sobel, motivated by the notion of allowing some uncertainty in the P(CC) considered both a conditional model [13] and an unconditional model [14] and developed optimal nested procedures for each of these models. In the conditional model, D is an upper bound on the number of defectives present and the calculations are all conditional on number of defectives present and the calculations are all conditional on this. In the unconditional model, D is viewed as the maximum number of defectives to look for and D is then determined so that the P(CC) is equal to or greater than a specified value P^* (close to one). In both cases when

$D = 1$ the procedure R_T agrees with the optimal nested procedure if N is reasonably small and p is very close to zero, (say, $Np < 1$).

If the model includes the assumption that there is at most $D = 1$ defective present among the N units, then the conditional expected number of tests is used to assess the procedure; for the procedure R_T this is

$$(21) \qquad E\{T|D = 1\} = \frac{q^N}{q^N + Npq^{N-1}}(1) + \frac{Npq^{N-1}}{q^N + Npq^{N-1}}\left(y + 1 + \frac{2R}{N}\right)$$

$$= \frac{q + p[N(y+1) + 2R]}{q + Np}$$

It has been noted above that for procedure R_T and this model the unconditional $P(CC)$-value given by (8) is still valid; the conditional probability of correctly classifying all units is of course equal to one under all these procedures. The conditional procedure proposed in [13] is an improvement on the procedure R_T since it does not restrict the strategy to a halving procedure. In actual application (partly because of the simple frequency interpretation it has) the unconditional approach is to be preferred, unless, of course, it is known that there is in fact at most $D = 1$ defective present.

The unconditional procedure R_{UD} in [14] would be carried out as follows. Suppose as before that $q = .9$ and $N = 15$ and we want to have a $P(CC) \geq P^*$, where $P^* = .94$ say. Then we find that the smallest integer D such that

$$(22) \qquad P\{X \leq D | N = 15, \; p = .1\} \geq P^*$$

is $D = 3$ by the use of any table of the cumulative binomial distribution. For comparison purposes it is desirable to randomize between $D = 2$ and $D = 3$ to make the probability in (22) exactly equal to .94 but for this illustration (or for practical usage) this can be avoided. The experimenter then proceeds to carry out the testing by using exactly the same strategy called R_1 that was extensively described in [7] as a function of N and $q = 1-p$. For example, it starts by examining 7 units. If these fail, it examines 3 of the 7. If the 7 units pass, it examines in the next test all 8 of the remaining units; the rest of the strategy is given in Table VA of [7]. The expected number of tests for $N = 15$, $q = .90$ and $D = 3$ is 6.8406.

The procedure stops as soon as 3 defective units are found or the units are found to be all good, whichever comes sooner. For this procedure $P(CC) = .9444$ when $N = 15$ and $q = .90$. This procedure also has an HLB and an ILB given in [14] and the values for $N = 15$, $q = .9$ are 6.7558 and 6.7025, respectively.

If a smaller P^*, say $P^* = .5$ had been chosen then one would find by (8) that $D = 1$ satisfies (22). The strategy remains the same except that the procedure now stops when 1 defective unit is found or when all the units are found to be good, whichever comes sooner. The expected number of tests for $N = 15$, $q = .90$ and $D = 1$ is 3.7606. Since this is exactly equal to the HLB for $D = 1$ in (10) above, it follows that the nested strategy for $D = 1$ in this case is an optimal procedure among all possible procedures, nested or otherwise. Indeed, this same result holds when $D = 1$ for any N and any q. For $D > 1$ one cannot show optimality in this manner but it is still conjectured to hold. It does not follow from this that $R_{U,D}$ is optimal among all nested procedures that satisfy the P^*-requirement (22), but this is also conjectured in [14] to be true for any N and any q.

A little care is necessary to avoid comparing procedures with different P(CC)-values. However, if one uses the same D-value in R_T and $R_{U,D}$ then the P(CC) is necessarily the same.

The optimal nested procedure for finding one defective (or classifying all the units as good) is the inconditional procedure $R_{U,1}$ defined in [14] without starting with any specified P^*, but just by letting $D = 1$.

For the case of an infinite population, i.e., $N = \infty$, the problem of finding a single defective was considered in [3]. The optimal nested procedure described there uses a mixed strategy (called R_1') that follows the procedure R_1 of [7] when a set is known to be defective (G-situation) and it has to be partitioned. Before that, in the so-called H-situation, this optimal procedure follws the strategy of another procedure R_2 (See Appendix A of [7]). For $N \to \infty$ this mixed strategy R_1' is an asymptotically optimal nested procedure. The advantage of letting $N \to \infty$ is that the resulting procedure that does not depend on N. On the other hand, the disadvantage is that for any fixed current value n of N the strategy will sooner or later call for a test batch of size x that is greater than n; from that point on one can no longer use the asymptotic procedure.

An <u>efficiency</u> rating (which is a function of q and N) can be defined for any procedure R by considering the inverse ratio of the expected number of tests for procedure R to the same quantity for the optimal nested procedure R_1'. The <u>asymptotic</u> efficiency (which is a function of q only) is the limit of this ratio as $N \to \infty$. Under this definition any nested halving procedure R that starts by testing all the N units has asymptotic efficiency zero for all q-values; this is becasue the value of $H_T(N|R)$ is approximately $\log_2 N$ which approaches ∞ as N gets large. On the other hand, the optimal result for finding one defective approaches a function of x (the initial batch size) that depends only on q and not on N. Hence the asymptotic efficiency is zero for any such procedure R. It should be noted that it is still possible for the efficiency of R for fixed q to be 100% for small or moderate values of N; the procedure R_T for $N = 15$ and $q = .99$ is a case in point as is indicated by Table 1 of [13]. Moreover, essentially the same argument shows that for the classical halving procedure R_0, which starts by testing $[N/2]$ units, is nested, and continues to partition by halving, the asymptotic $(N \to \infty)$ efficiency is again zero.

To discuss the efficiency of procedure $R_{1,D}$ defined by recursive formulae in [13], it is necessary to take into account the fact that it is assumed in the model that there are at most D defectives among the N units. Thus conditional probabilities are appropriate here. Since the HLB is actually attained for $D = 1$, this procedure R_{11} is 100% efficient (when $D = 1$) for all N and all q under that model; hence for $D = 1$ we also have 100% asymptotic efficiency over all q-values under that model. For $D \geq 2$ this procedure $R_{1,D}$ is conjectured to be efficient for all N and all q, but this was not proved.

In a companion paper [14] the procedure $R_{U,D}$ for the corresponding unconditional problem allows us to look for at most D defectives without assuming that there are at most D defectives among the N units. Under this unconditional model we can make fair comparisons of the efficiency of $R_{U,1}$, R_T and R_0 all of which are appropriate for problem 2 and do not assume (as part of the model) that there is at most $D = 1$ defective unit among all the N units. This procedure $R_{U,1}$ described in [14] uses exactly the same strategy as the procedure R_1 in [7] for all D, N and q-values. Moreover, for $D = 1$ it appears that the HLB is attained for all N and all q, so that procedure $R_{U.1}$ is 100% efficient for all N and all q-values and this property is conjectured to hold for all values of D. Thus procedure $R_{U,1}$ serves as a natural base for computing the efficiency of any other procedure for problem 2. Extensive computations for finite N have not been made. However, as noted earlier, the asymptotic efficiency is zero for both R_T and R_0 and indeed for any procedure that is nested, tests a fixed (positive) proportion of all the N units at the outset, and also tests another fixed (positive) proportion of all the units in a defective batch. The halving procedures under discussion (like R_T and R_0) clearly have these properties.

Problem 3: Finding all the Defectives.

Two other halving procedures, R_4 and R_5, are introduced in Appendix C of [7] but these are expressley set up to find all the defectives in the sample of N units. Although one cannot compare these directly with R_T or $R_{U,D}$, it should be of considerable interest to the user to know how much more testing is needed on the average to raise his P(CC) to one. One of these procedures, R_4, allows the experimenter to combine independent binomial sets (in the current state of knowledge) and hence it is called a recombination procedure. Under this procedure one only needs to keep track of at most two batches of unclassified units. Under the other procedure, R_5, recombination is not allowed and the experimenter has to work with more and more batches of units as the testing proceeds. The latter, R_5, is less efficient and should be used only when the application forbids any recombination. Explicit formulas and numerical values for the expected number of tests required under procedures R_4 and R_5 for $N \le 12$ are given in Appendix C of [7].

To enlarge the table for procedure R_4 in equation (151) of [7] the polynomials for the expected number of tests for $N = 13, 14, 15$ and 16 have been derived; from equations (149) and (150) of [7] for the halving procedure R_4 these results are

$$H_4(13) = 41 - 32q - 5q^2 - q^3 + 2q^4 - 5q^5 + q^6 + 3q^7 - 2q^8 - q^9, \tag{23}$$

$$H_4(14) = 45 - 35q - 6q^2 - q^3 + 3q^4 - 6q^5 + q^6 + 3q^7 - 2q^8 - q^9, \tag{24}$$

$$H_4(15) = 49 - 38q - 7q^2 - q^3 + 3q^4 - 6q^5 + q^6 + 3q^7 - q^8 - 2q^9, \tag{25}$$

$$H_4(16) = 54 - 42q - 8q^2 - q^3 + 3q^4 - 6q^5 + q^6 + 3q^7 - q^8 - 2q^9 . \tag{26}$$

For $N = 15$ and $q = .9$ equation (25) gives us 7.5874 as the expected number of tests to classify all the units using the halving procedure R_4. In contrast, the procedure R_1 for the same N and q requires 7.213 tests on the average; the latter appears in Table VA of [7]. The HLB for this problem is 7.085 and the ILB is 7.035; these can be found in Table 2 A of [9]. Although the HLB is not met, the procedure R_1 is known to be optimal among all nested procedures. For further contrast, the original procedure R_D^* of Dorfman in [1] breaks up the $N = 15$ units into 3 batches of size 4 and 1 batch of size 3 when $q = .90$. For each batch, the first test is on the whole batch and if that fails each unit is tested separately. No inference is used if a batch of 4 fails and the first 3 of these units are found to be good, i.e., the 4th unit has to be tested also; a modification of Dorfman's procedure that uses inference was introduced in [5] but we omit it in the present discussion. The expected number of tests under R_D^* for $N = 15$ and $q = .90$ is

$$(27) \qquad E\{T|R_D^*\} = 3[5 - 4q^4] + 4 - 3q^3 = 19 - 3q^3 - 12q^4 = 8.9398 .$$

Another modification R_S of Dorfman's procedure is due to Sterrett [15], which also fails to bring in the use of inference when it is applicable. In this modification defective batches are tested one-at-a-time only until a defective unit is found. Then the remainder of that batch is tested and either it is passed in 1 test or we again test units individually until a defective unit is found, etc. In computing the result for the Sterrett procedure we have added the use of inference so that the results are slightly better for this modification R_S^* than for the procedure R_S as originally proposed. Recombination of binomial units is not used in R_S or in R_S^*. The starting value n is of course the optimal value for the modified procedure R_S^* based on the known value of q. For $q = .90$ we use groups of size $n = 5$ for R_S^* (and also for R_S), so that for $N = 15$, we have three such groups. The expected number of tests under R_S^* for each group of size n and for any q is given by

$$(28) \qquad H_S(n) = n + 1 + (n-2)p - \frac{q}{p}(1-q^n) ,$$

but we shall not derive this result here. (See also the discussion about this in [4] and [9]). It should be noted that no exact expression like (28) is given in [15] for the procedure R_S. The expected number of tests under R_S^* for $N = 15$ and $q = .90$ is therefore

$$(29) \qquad E(T|R_S^*) = 3H_S(5) = 7.8432 ,$$

which is an improvement over the result for procedure R_D^*. Thus, the halving procedure R_4 is an improvement over both R_D^* and R_S^* and the optimal nested procedure R_1 is a substantial improvement over both R_D^* and R_S^*.

These improvements could be translated into an asymptotic efficiency rating for any procedure R (as a function of q) by considering for any N the inverse ratio of the expected numbers of tests for the procedure R and the optimal nested procedure and finding the limit as $N \to \infty$ of this ratio. This has been done in [4] for a number of procedures; the foundation for this work is in [10]. [All efficiencies in the ensuing discussion are meant to be asymptotic $(N \to \infty)$ and this adjective will be dropped at times when there is no confusion. To be specific the efficiency is usually computed at $q = .99$.] Thus, for $q = .99$ it was found that the asymptotic efficiency of the Dorfman procedure R_D^* is only 41.44% and that of R_S^* is somewhat higher, 54.44%. By the same definition the procedure R_1 has 100% efficiency with respect to the class of nested procedures and is conjectured to be 100% efficient among all procedures.

Another procedure R_2 based on information theory concepts is introduced in Appendix A of [7]. It represents a limit $(N \to \infty)$ of the procedure R_1 (R_1 is uniformly better than R_2 for all q and all N) and the recursion for R_2 depends on q, but not on N. The asymptotic efficiency of R_2 is shown in [4] to be over 99% uniformly in q among all procedures, nested or otherwise.

The asymptotic efficiency of procedure R_4 was not considered in [4] and this has not yet been evaluated. For q close to 1, say $q = .99$, it should prove to be highly efficient. However, for $q < 1/2$ it has been noted in Table II A of [7] that both procedures, R_4 and R_5, are highly inefficient. In fact, R_4 is even worse than R_5 and both efficiencies apparently approach zero as $q \to 0$. One could argue that no group-testing procedure would be used if it was known that $q < 1/2$. but the halving procedures are also useful when q is unknown and these efficiencies still apply.

There is a halving procedure, R_{04}, defined on page 139 of [10] that is highly efficient for all q. Here halving is used only in the so-called G-situation when a set is known to be defective; in the so-called H-situation an optimal batch size x is found by searching for that integer x that minimizes the expression

$$(30) \qquad W(q|R_{04}) = \left[\frac{1 + p\, F^{*}(x)}{1 - q^{x}} \right] p \ ,$$

where $F^*(x)$, which depends on q, is the same function that appears in (2) and (3) above. For $q < 1/2$ it is easily found that $x = 1$ so that the efficiency is 100% for $q < 1/2$ and, in fact, this holds for $q < (\sqrt{5}-1)/2 = .618...$ exactly as for the optimal procedure. The q-intervals for $x = 2$, $x = 3$ and $x = 4$ in the H-situation under procedure R_{04} are exactly the same as under the asymptotically optimal procedure (See R_{21} and R_{01} both defined in [10]). Hence up to $q = .8813$ (and also for other higher values of q) the procedure R_{04} is 100% efficient. In the H-situation procedure R_{21} (and R_{01}) uses $x = 5$ for $.8567 \le q \le .8813$, $x = 6$ for $.8813 \le q \le .8987$, and $x = 7$ for $.8987 \le q \le .9116$; in general it uses x for q in the interval bounded on the left by the root of $1-q^{x-1} - q^x = 0$ and on the right by the root of $1-q^x - q^{x+1} = 0$. Corresponding to this, the procedure R_{04} uses $x = 5$ for $.8567 \le q \le .8897$, and uses $x = 7$ for $.8897 \le q \le .9116$, so that $x = 6$ is never used in an H-situation under R_{04} for any q; the value $q = .8897$ is the root of $1-q^5-q^7 = 0$. In the G-situation with these small values of x, the procedure R_{04} is the same as that of R_{21} (or R_{01}) for all $x \le 9$, with the exception of $x = 6$. It follows that the procedure R_{04} is also 100% efficient for $.8987 \le q \le .9296$, since R_{21} uses $x = 7$ for $.8987 \le q \le .9116$, $x = 8$ for $.9116 \le q \le .9216$ and $x = 9$ for $.9216 \le q \le .9296$. From these remarks one can expect the minimum efficiency of R_{04} (over all $q < .93$) to occur near $q = .89$, at the center of the interval above (.8813, 8987), where it disagrees with R_{21}. Calculating the value of $W(q|R_{04})$ in (28), one obtains .5055 which represents the (asymptotic or long term) expected number of tests per unit classified when $q = .89$

and $N = \infty$. The corresponding quantity for procedure R_{21} (or R_{01}) is .5032. The ratio 5032/5055 = .995 or 99.5% is therefore estimated to be the minimal efficiency of procedure R_{04} over all values of $q < .93$. For $q = .95$ and .99 the efficiency of R_{04} is even higher than 99.5% (See Table 1 of [10]), but 99.5% is still not the minimum asymptotic efficiency over all values of q. As q approaches one and the value of x (for the H-situation) increases, the procedure R_{04} continues to prefer powers of two for x and to avoid numbers of the form $3 \cdot 2^y$ (like 6) which are half-way between powers of two. The latter values of x correspond to values of q at which the efficiency of R_{04} "hits a valley". Thus at $q = .945$ the value of x for R_{21} is 12 but procedure R_{04} uses $x = 15$. The efficiency of R_{04} at $q = .945$ is 99.4%; this is slightly lower than the value at $q = .89$, but still over 99%. At $q = .9715$ the value of x for R_{21} is 24 but procedure R_{04} uses $x = 31$; the efficiency of R_{04} at $q = .9715$ is 99.2%. At $q = .9857$ the value of x for R_{21} is 48 but procedure R_{04} uses $x = 63$; the efficiency of R_{04} at $q = .9857$ is again 99.2%. For q values $> .9857$ the low efficiency values (when x is of the form $3 \cdot 2^y$) finally start to increase toward one and it appears as if the efficiency values of 99.2% is a global minimum over all q-values.

In summary, if one regards procedure R_{04} defined in [10] as a halving procedure (it uses halving only in the G-situation), then it has been shown by actual construction that there exists a halving procedure for finding all the defectives which has very high asymptotic efficiency (it appears to be > 99%) uniformly for all values of q. However, it should be pointed out that procedure R_{21} (or R_{01}) is uniformly better than R_{04} and, as long as the instructions and tables are available for both, one could argue that there is little reason for preferring R_{04} to R_{21}.

BIBLIOGRAPHY

[1] Dorfman, R. (1943). The detection of defective members of large populations. Annals of Math. Stat. 14, 436-440.

[2] Huffman, D. A. (1952). A method for the construction of minimum redundancy codes. Proc. I. R. E. 40, 1098.

[3] Kumar, S. and Sobel, M. (1971). Finding a single defective in binomial group-testing. JASA 66, 824-828.

[4] Kumar, S. and Sobel, M. Group-testing with at most c tests for finite c and $c \to \infty$. University of Minnesota Technical Report No. 146. (submitted for publication).

[5] Lee, J. K. and Sobel, M. (1972). Dorfman and R_1-type procedures for a generalized group-testing problem. Mathematical Bio-Sciences 15, 317-340.

[6] Sandelius, M. (1961). On an optimal search procedure. Amer. Math. Monthly 68, 133-134.

[7] Sobel, M. and Groll, P. A. (1959). Group-testing to eliminate efficiently all defectives in a binomial sample. Bell System Tech. Journal 38, 1179-1252.

[8] Sobel, M. and Kumar, S. (1973). Asymptotically Optimal Bayes Solution in Group-testing for an Infinite Population. (To appear in Symposium on "Linear Statistical Inference" Ed. by J. N. Srivastava).

[9] Sobel, M. and Groll, P. (1966). Binomial group-testing with an unknown proportion of defectives. Technometrics 8, 631-656.

[10] Sobel, M. (1960). Group-testing to classify all defectives in a binomial sample. A contribution in Information and Decision Processes. Ed. R. E. Machol, McGraw-Hill, New York, pp. 127-161.

[11] Sobel, M. (1967). Optimal group-testing. Proceedings of the Colloquim on Information Theory organized by the Bolyai Mathematical Society, Debrecen (Hungary), 411-488.

[12] Sobel, M. (1968). Binomial and Hypergeometric group-testing. Studia Scientiarum Mathematicarum Hungarica 3, 19-42.

[13] Sobel, M. (1973). Group-testing with at most D defectives: Binomial and Hypergeometric Models. Univ. of Minnesota School of Statistics Technical Report No. 204.

[14] Sobel, M. (1973). Group-testing with Preassigned Confidence Levels for a correct classification of all units. Univ. of Minnesota School of Statistics Technical Report No. 223.

[15] Sterrett, A. (1957). On the detection of defective members of large populations. Ann. Math. Statist. 28 1033 - 1036.

[16] Thomas, J., Pasternack, B. S., Vacirca, S. J. and Thompson, D. L. (1973). Application of Group-testing Procedures in Radiological Health. Health Physics 25, 259-266.

[17] Ungar, P. (1960). Cut-off point for group-testing. Comm. on Pure and Applied Math. 13, 49-54.

[18] Zimmerman, S. (1959). An optimal search procedure. Amer. Math. Monthly 66, 690-693.